

AD-A037 365

NORTHROP NORTRONICS ANAHFIM CALIF SYSTEMS SUPPORT DEPT
A TOPOLOGICAL FOUNDATION FOR SELF-ORGANIZATION, (U)
NOV 63 R I SCIBOR-MARCHOCKI

F/G 9/4

UNCLASSIFIED

NSS-2828

NL

1 OF 1
AD
A037365

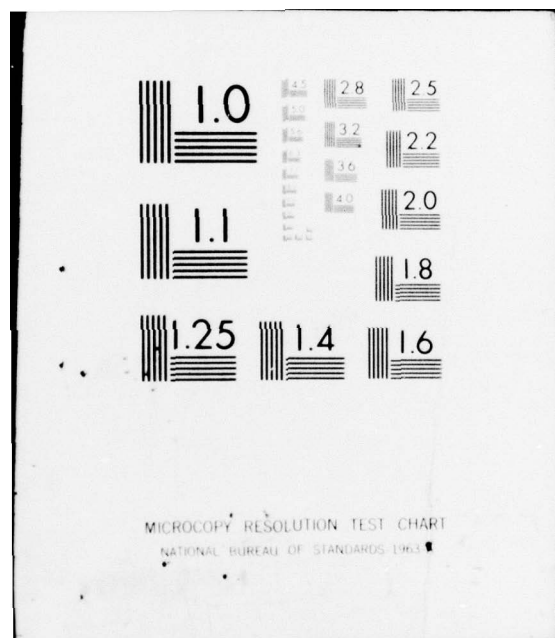


END

DATE

FILMED

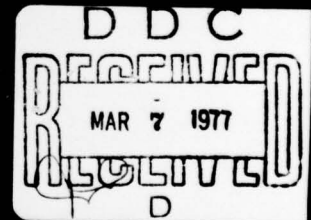
4-77



ADA 037365

NORTRONICS

A DIVISION OF NORTHROP CORPORATION



NORTHROP

A TOPOLOGICAL FOUNDATION
FOR SELF-ORGANIZATION

DISTRIBUTION STATEMENT A

Approved for public release;
Distribution Unlimited

ADDITIONAL TO	
TYPE	Write Section <input checked="" type="checkbox"/>
SEC	Diff Section <input type="checkbox"/>
STANDARD	<input type="checkbox"/>
Per H. on file	
CLASSIFICATION	
CLASSIFICATION CODES	
SPECIAL	
A	

MOST Project - 7

14 NSS Report No. 2828
1981-5156

6
A TOPOLOGICAL FOUNDATION
FOR SELF-ORGANIZATION

Presented 14 November 1963 at the
First Pasadena Institutional Symposium
on Self-Organizing Systems by
Romuald I. Scibor-Marchocki

10

410 114

DDC
RECEIVED
MAR 7 1977
D

NORTHROP NORTRONICS

SYSTEMS SUPPORT DEPARTMENT
500 East Orangethorpe Avenue
Anaheim, California

11 14 November 1963

DISTRIBUTION STATEMENT A
Approved for public release;
Distribution Unlimited

12 8p.

lpg

A Topological Foundation for Self-Organization
Romuald I. Ścibor-Marchocki

It is shown that by the use of Information Theory, any metrizable topology may be metrized as an orthogonal Euclidean space (with a random Gaussian probability distribution) times a denumerable random cartesian product of irreducible (wrt direct product) denumerable groups. The necessary algorithm to accomplish this metrization from a statistical basis is presented. If such a basis is unavailable, a certain nilpotent projection operator has to be used instead, as is shown in detail in the companion paper. This operator possesses self-organizing features.

INTRODUCTION

In the companion article¹ we will define a self-organizing system as one which, after observing the input and output of an unknown phenomenon (transfer relation), organizes itself into a simulation of the unknown phenomenon.

Within the mathematical model, the aforementioned phenomenon may be represented as a topological space thus omitting for the moment the (arbitrary) designation of input and output which, as will be shown, bears on the question of uniqueness. Hence, for the purpose of this paper which emphasizes the mathematical foundation, an intelligent device is taken as one which carries out the task of studying a space and describing it.

In keeping with the policy that one should not ask someone (or something) else to do a task that he could not do himself (at least in principle), let us consider how we would approach such a problem.

In the first place, we have to select the space in which the problem is to be set. The most general space that we feel capable of tackling is a metrizable topology. On the other hand, anything less general would be unnecessarily restrictive. Thus, we choose a metrizable topological space.

As soon as we have made this choice, we regret it. In order to improve the situation somewhat, we show that there is no (additional) loss of generality in using an orthogonal Euclidean space times² a denumerable random cartesian product of irreducible (wrt direct product) denumerable groups.

This paper provides a survey of the problem and a method for solving it which is conceptually clear but not very practical. The companion paper¹ provides a practical method for solving this problem by means of the successive use of a certain nilpotent projection operator.

METRIZATION

We start with a metrizable topological space. There are many equivalent axiomatizations of a metrizable topology; eg, see Kelley³. Perhaps the easiest way to visualize a metrizable topology is to consider that one was given a metric space but that he lost his notes in which the exact form of the metric was written down. Thus one knows that he can do everything that he could in a metric space, if only he can figure out how.

The "figuring out how" is by no means trivial. Here, it will be assumed that a cumulative probability distribution has been obtained on the space by one of the standard methods; bird in cage⁴, Munroe⁵ I, Munroe⁵ II, ordering (see Halmos⁶ or Kelley³), or Kelley³. This cumulative probability distribution is a function on X onto the interval $[0, 1]$ of real numbers. The inverse of this function, which exists by the Radon Nikodym theorem, provides a mapping from the real interval onto the non-trivial portion of X . This mapping induces all of the pleasant properties of the real numbers on the space X : topological, metric, and ordering.

Actually, it turns out that, especially if the dimensionality of the space is greater than one, the foregoing procedure not only provides one metrization, but many. Indeed, this lack of uniqueness is what makes the procedure exceedingly difficult. Only by imposing some additional conditions that result in the existence of a unique solution, does the problem become tractable.

We choose to impose the additional condition that the resulting metric space be a Euclidean geometry with a rectangular coordinate system. Even this always does not yield uniqueness, but we will show the additional restriction that will guarantee uniqueness after the necessary language is developed. Since all metrizations of a given metrizable topology are isomorphic, in the quotient class the orthogonal Euclidean geometry serves the purpose of being a convenient representative of the unique element resulting from a given metrizable topology.

Furthermore, the same comment applies to the use of a Gaussian distribution as the probability distribution on this orthogonal Euclidean geometry. Namely, the random Gaussian distribution on an orthogonal Euclidean geometry is a convenient representative member of the equivalence class which maps into one element (stochastic space) of the quotient class.

Information Theory

Now, we will show that Information Theory provides the language necessary to describe the metrization procedure in detail.

It is possible to introduce Information Theory axiomatically by a suitable generalization of the axioms⁷ in Feinstein⁸. But to simplify the discussion here, we will use the less elegant but equivalent method of defining certain definite integrals. The probability density distribution p is defined from the cumulative probability distribution ρ by

$$(1) \quad \rho(X') = \int_{X' \text{ measurable} \subset X} p(x) dx, \quad ,$$

Then the information rate H is defined as

$$(2) \quad H(X) = - \int_X p(x) \ln \kappa p(x) dx, \quad ,$$

where κ has (carries) the units of X . Finally, the channel rate R is defined as

$$(3) \quad R(\bigotimes_I X_i) = \sum_I H(X_i) - H(X), \quad ,$$

where X is the denumerable⁹ cartesian product space

$$(4) \quad X = \bigotimes_I X_i.$$

Next, we define the angle θ

$$(5) \quad |\theta(\bigcirc_I X_1)| = \sin^{-1} e^{-R(\bigcirc_I X_1)}$$

and the norm

$$(6) \quad |X| = \kappa (2\pi e)^{-1/2} e^{H(X)}.$$

Now, if¹⁰ a statistically independent basis; ie, one for which

$$(7) \quad R(\bigcirc_I X_1) \stackrel{\kappa}{=} \text{constant},$$

can be provided in terms of one dimensional components; ie, none of them can be decomposed further, then it is just the usual problem of diagonalization of a symmetric matrix by means of a congruence transformation to provide an orthogonal coordinate system. Furthermore, for uniqueness, we arrange the spectrum in decreasing order. Then, by means of the Radon Nikodym theorem applied to each of these one-dimensional axes, the probability distribution may be made; eg, Gaussian, if desired. Thus, we obtain the promised orthogonal Euclidean space.

Channel

At this time we can state the remaining additional condition required that a decomposition be unique. The index space I has to be partitioned into exactly two parts, say I' and I'' ; ie,

$$(8) \quad \begin{aligned} I' \cup I'' &= I \\ I' \cap I'' &= \emptyset, \end{aligned}$$

such that

$$(9) \quad \dim(X') = \dim(X''),$$

where

$$(10) \quad \begin{aligned} X' &= \bigotimes_{I'} X_1 \\ X'' &= \bigotimes_{I''} X_1. \end{aligned}$$

(If $\dim(X)$ is odd, then we have to cheat a little by putting in an extra random dummy dimension.) And then the decomposition of the space

$$(11) \quad X = \bigotimes_I X_1$$

has to be carried out so that this partitioning is preserved. Since this partitioning is arbitrary (as far as the mathematics is concerned), it is obvious that a space which is not partitioned will have many (equivalent) decompositions. On the other hand, if the partitioning is into more than two parts, then the existence of a decomposition is not guaranteed.

A slight penalty has to be paid for the use of this partitioning, namely: instead of eventually obtaining a random cartesian product of one-dimensional spaces, we obtain an extended channel (with random input) of single-dimensional channels. It is obvious that if we were to drop the partitioning temporarily, each such single-dimensional channel would be further decomposed into two random components. This decomposition is not unique. But one of these equivalent decompositions is particularly convenient; namely, that decomposition where we take the component out of the original X' and that which is random to it, say V . This V (as well as the cartesian product of all such V 's, which of necessity are random) is called the linearly additive noise. The name "linearly additive" is justified because it is just the statistical concept isomorphic to the linear addition of vectors in orthogonal Euclidean geometry. (The proof of this last statement is not completed as yet.)

Denumerable Space

The procedure for this decomposition was worded to de-emphasize the possible presence of a denumerable (component of the) space. Such a component may be given outright; otherwise, it results iff the space was not simply connected. Any denumerable space is zero dimensional, as may be verified easily from the full information theoretic definition of dimensionality.

The obvious way of disposing of a denumerable space is to use the conventional mapping⁵ that converts a Stieltjes to a Lebesgue integral, using fixed length segments. (It can be shown that H is invariant under such a mapping.) Unfortunately, while this mapping followed by a repetition of the preceding procedure will always solve a given problem (no new¹¹ denumerable component need be generated on the second pass), little insight is provided into the structure of the resulting space. On the other hand, because channels under cascading constitute a group, any such denumerable space is a representation of a denumerable group.

SUMMARY

In summary, the original metrizable topological space was decomposed into an orthogonal Euclidean space times² a denumerable random cartesian product of irreducible (wrt direct product) denumerable groups. Thus, since any individual component of a random cartesian product may be studied independently of the others, all that one needs to study is; (1) a Gaussian distribution on a single real axis and (2) the irreducible denumerable groups.

Finally, it should be emphasized that there are only these two ways of decomposing a metrizable topology; (1) if a (statistical) basis is given, use the diagonalization of a symmetric matrix algorithm described earlier (and given in detail in the three channels in cascade problem), and (2) otherwise use a suitable network of the NPO's with $n_0 = 1$. Of course, any hybrid of these two methods may be employed as well.

¹Peter A. Kleyn, Conceptual Design of Self-Organizing Machines, (Northrop Nortronics, Anaheim, California, 14 November 1963), NSS Report No. 2832.

²Random cartesian product.

³John L. Kelley, General Topology, (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1955).

⁴Willis W. Harman, Principles of the Statistical Theory of Communication, (McGraw-Hill Book Company, Inc., New York, New York, 1963).

⁵M. E. Munroe, Introduction to Measure and Integration, (Addison-Wesley Publishing Company, Inc., Cambridge, Massachusetts, 1953).

⁶Paul R. Halmos, Measure Theory, (D. Van Nostrand Company, Inc., Princeton, New Jersey, 1950).

⁷Feinstein⁸ uses his axioms only in a finite space X ; ie, $\text{card}(X) < \aleph_0$.

⁸Amiel Feinstein, Foundations of Information Theory, (McGraw-Hill Book Company, Inc., New York, New York, 1958).

⁹If I is infinite, certain precautions have to be exercised.

¹⁰This "if" is the catch that makes all methods of metrization of a space of dimensionality higher than one impractical, except the method of successive projections upon unit spheres centered at the center of gravity. The method of using that nilpotent projection operator is described in the companion paper¹.

¹¹Only non-cyclic irreducible (wrt direct product) denumerable group components of the old denumerable space will remain.

FILM
4